

AD-A192 837 GENERIC CUEING IN IMAGE UNDERSTANDING(II) ROYAL SIGNALS 1/1

AND SIGNAL ESTABLISHMENT, WATFORD (ENGLAND)

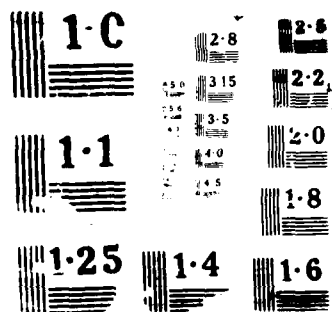
P. FREEMAN ET AL. DEC 87 ASKE-MEMO-4189 DRIC-88-105235

UNCLASSIFIED

F/C 12/8

ML





AD-A192 037

ROYAL SIGNALS AND RADAR ESTABLISHMENT

Memorandum 4109

TITLE: GENERIC CUEING IN IMAGE UNDERSTANDING  
AUTHORS: P. Fretwell, D.A. Bayliss, C.J. Radford, R.W. Series  
DATE: December 1987

SUMMARY

We report an attempt to derive simple low level features which may readily be extracted from an image. The features may be used to provide a classification of objects, so acting as a cue to aid further recognition.

It is shown that such an approach may find applications in the early stages of image analysis for the classification of objects in an open world situation. The method is illustrated by application of classification rules based on an estimate of line wigglyness (fractal dimension) and analysis of the directional edge statistics.

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Copyright  
C  
Controller HMSO London

1987

## CONTENTS

1. INTRODUCTION	1
2. RELATED TECHNIQUES	3
3. GENERIC IMAGE CUEING	4
3.1 Cueing with edge wigglyness	5
3.2 Cueing with edge direction	7
4. DISCUSSION AND FUTURE WORK	8
5. CONCLUSION	9
REFERENCES	

# GENERIC CUEING IN IMAGE UNDERSTANDING

P. Fretwell, D.A. Bayliss, C.J. Radford, R.W. Series,

RSRE,

St Andrews Road,

Malvern,

Worcs,

WR14 3PS

## Abstract

We report an attempt to derive simple low level features which may readily be extracted from an image. The features may be used to provide a classification of objects, so acting as a cue to aid further recognition.

It is shown that such an approach may find applications in the early stages of image analysis for the classification of objects in an open world situation. The method is illustrated by application of classification rules based on an estimate of line wigglyness (fractal dimension) and analysis of the directional edgel statistics.

## 1 Introduction

In image understanding a cue may be defined as a link between one level of object description to a higher level via measurement at the lower level. The ability to reason over uncertain and possibly contradictory inferences derived from this cue analysis will be the subject of a future publication.

In this paper we report the use of a combination of generic features to form specific object cues for the classification of an open world scene. It is shown that

taken individually the individual cues provide limited discrimination but certain combinations of them can provide an effective object cue. An edgel representation of the scene which has been edge extracted from a grey level image by the application of a directional edge operator is used as the basis for the study. The edgel representation used has a five bit edge strength component and a three bit (eight directions) directional component.

A simple, computationally inexpensive cueing method that links directional edge information to some as yet unspecified higher level model is proposed. The domain is real images set in natural and semi-natural outdoor environments. This choice of domain poses many difficulties not found in closed world problems such as are encountered in certain industrial part recognition problems in highly structured and controlled environments. One of the many difficulties in the open world scenario is that the background cannot be controlled and this can cause problems when trying to match an object model against the image. Another difficulty the open world can share with the closed world is that of object variability. This may be found at three levels. Most simply, a given object, from a given viewpoint, may take a variety of appearances depending on the illumination. Next, the image will vary with viewpoint. Finally, in most open world situations, there is no unique three dimensional specification of the object and we require to classify objects in terms of their generic structure.

Our approach is not to use an exact representation of the three dimensional shape of the object. Rather, the idea is to use generic image properties of an object class to approximately identify the location and identity of the object in the image. We suggest that such an approach may find applications in the early stages of image analysis for the classification of objects in an open world situation. This will of course be followed by verification of the exact identity and position of the object under scrutiny. A verification method could use a spatial model based

approach, see for instance Brooks 1981, Lowe 1985, 1987 and Brisdon 1987 and the related paper Sullivan 1987.

## **2 Related Techniques**

The cue rules in the method presented in this paper are based on an estimate of wigglyness of edge features and analysis of directional edgel distribution. Details of this approach appear in the next section. Related work includes that of Knoll and Jain 1986 and more recently Wallace 1987.

Knoll and Jain use binary images of three dimensional flat objects. Using what amount to two dimensional objects overcomes the problems of viewing variability. Their work is close in philosophy to the method proposed in this paper in that they identify common features that objects in a particular class possess. Therefore, identification does not rely on unique features. This means it can cope with some obscuration. The common features in the method are lengths of certain parts of the objects boundaries. This makes the method dependent on some measure of scale. Our method is virtually scale independent and does not rely on the measurement of image edge segments. The breaking up of continuous object lines by edge operators into smaller segments is a major problem when analysing edge maps for image understanding. This non robustness derives from the possible shortfall in the edge operators performance and more importantly from the fact that a pixel value is a product of so many factors. It depends on the objects reflectance, the light source, the orientation of the object with respect to the light source and so on. The method of Knoll and Jain is particularly affected by the output of the edge operator. Our method has a certain robustness to breaking up of the edges.

Wallace 1987 concentrates on the recognition of flat industrial parts based on the identification of binary shape cues. These cues are pairs of line segments in a particular relationship such as roughly joined at right angles. They claim that the



method should be robust to partial obscuration because not all the object's boundary is used. However, the model used is quite specific to a particular object and is prone to the problems of object variability. The variability caused by different viewing angles will cause the number of cues for the object to become very large indeed. The technique is used to recognise flat industrial parts. The background is uncluttered and the parts appear alone or in twos in the image. It is not clear what the outcome would be if the background was to become cluttered or more parts were added to the scene. It is probable that the number of combinations of cues necessary to identify the object would become exponentially large as the background and other parts in the image contributed more and more false cues.

In contrast our method uses generic properties of the object image. These image properties are not necessarily unique to the object but they do give an indication of the presence of the object. Various combinations of these image properties can serve to strengthen the belief that a particular object is at a particular position. This means that background properties will influence the cueing but will not necessarily mean that as the background complexity increases, the method will take an exponentially increasing amount of time.

A related approach uses a region segmentation together with contextual reasoning to label the regions of the segmentation. (Golden, Fullwood and Hyde 1987, Morton 1987). The output of these methods is a set of bounded regions in the two dimensional image in which an object of interest is thought to be located, along with an initial view point hypothesis.

### **3 Generic Image Cues**

A series of cues are desired that can roughly indicate the presence of objects in a large object class. In other words the cue is designed to use only those attributes that all members of the class share. Thus it trades in accuracy for generality. The



Figure 1: Typical Test Set Car Image



Figure 2: Edgel Map of Typical House Image

purpose of this work is to propose several such cues and investigate how they may be combined to produce a more accurate cueing mechanism.

### 3.1 Cueing with Edge Wigglyness

Figure 1 is a typical image taken from a test set of 60 images used in this work. Figure 2 is an edgel map of a typical image containing a building taken from the same test set. It was recognised, not surprisingly perhaps, that the buildings and cars had more linear features (mainly, straight lines) while the bushes and trees have more high frequency in their edge structure.

To investigate the applicability of this classification we used an edge extraction operator due to the work of Radford (to be published). This operates first by extracting an edge map from the grey level image using an adaptive directional Sobel operator. This provides a series of edge segments, each of which has information about local direction coded into the output. A second stage has the effect of tracking along the lines and measuring the total change in line direction as a function of line length. This provides a measure of the "wigglyness" of the line which is approximately independent of scale. Wigglyness has an inverse relationship to fractal dimension. Note that the code takes no precautions to detect a wiggly line joined onto a straight line and so is potentially capable of significant enhancement.

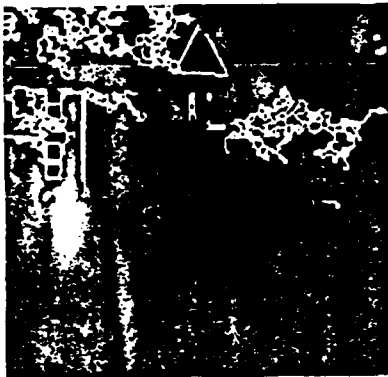


Figure 3: Wiggly Edgels



Figure 4: Non-Wiggly Edgels

Figure 3 shows those edgels that have been classified as "wiggly" while the image in Figure 4 shows only those edges classified as "non-wiggly" by the filter. Note that most of the edge features of the building and car are classified as smooth while much of the detail of the vegetation is classified as wiggly although not exclusively so as noted above. These are primarily due to the joining of edges of differing types. There are various methods which might be used to overcome this problem, for example a statistical analysis of the fractal dimension along a line might be used to look for changes in line character in a manner analogous to the DSRM algorithms used in some region growing algorithms (Godden, Fullwood and Hyde 1987).



Figure 5: Horizontal Edgels



Figure 6: Vertical Edgels

### 3.2 Cueing using Edgel Direction

Analysis of test images after application of the edge operator and wiggle filter revealed that the motor cars and the buildings in the images were responsible for many of the long straight horizontal lines. To study the potential value of the directional edge information, maps were produced of the vertical and horizontal edge components in a variety of images. As the edge operator produces output giving directional information as a three bit output, the maps show edgels within plus or minus 22.5 deg of horizontal or vertical directions

Figure 5 shows the edgels in the typical car image that have horizontal orientation (to within plus or minus 22.5 degrees) while Figure 6 shows the edgels in the typical house image that have vertical orientation. It was noted that the bushes and trees made a significant contribution to the horizontal edgels. The building regions are responsible for long vertical lines while the car and shrubbery regions are responsible for many of the short vertical lines. These findings form the basis of a simple cue for potential cars or buildings within the test data.

The ratio of the number of horizontal edgels to the number of vertical edgels was calculated for the regions that contained cars and buildings. It was found that the ratio for the two different regions was different and also had a value that was fairly constant (within a range of values) over the majority of the test set. This was in spite of the cars having different scale and orientations (within a horizontal plane) within the image. Thus the ratio could be said to have reasonable scale invariance. It was found that if the edgels corresponding to the vegetation were removed the ratio for the car and building regions were more consistent.

From this observation an algorithm was devised that used the vegetation removal procedure and the ratio test for regions of an image to discriminate between car like and non-car like and house like and non-house like.

The algorithm first removes the majority of the edgels corresponding to the

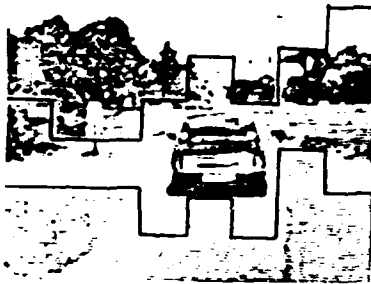


Figure 7: Car Cue



Figure 8: Building Cues

vegetation as possible. The parameter governing the edgel operator as well as the vegetation removal is achieved by fixed thresholds. The resulting edgel image is then divided into 64 regular regions. The ratio of horizontal edgels to vertical edgels within each region is calculated. If the region has a ratio that lies outside  $[1.0, 60.0]$  then the region is said not to contain a car and similarly if the ratio lies outside  $[0.2, 10.0]$  then the region is said not to contain a building. Finally the largest contiguous set of regions with the appropriate ratio is selected as the cue region.

The algorithm is demonstrated in Figure 7 and Figure 8. The dark regions in Figure 7 correspond to those regions that have not got a ratio that lies inside the car ratio. Figure 8 shows the algorithm for the case of the building. It can be seen that the majority of the indicated region in Figure 7 contain most of the car. Similarly for the case of the building.

#### 4 Discussion and Future Work

Cueing based on the ratio of horizontal to vertical edgels within an image is clearly sensitive to variations in the edgel operator output. This in turn is sensitive to changes in grey level caused by lighting, shadows, reflectance and so forth. This

potential problem maybe overcome by performing the cueing using several edgel maps of different sensitivities and combining these to obtain a ratio. This approach could allow the cueing method to be independent of hand crafted parameters in the edgel operator.

The cueing is based on analysis of 64 rectangular grid regions in the edgel domain. Unless the object fills a whole number of regions there will be some regions that contain only a small part of the object. This may cause the cueing to disregard those parts of the image. The cueing could be more accurate if there were fewer regions that only contained a small part of the object. This could possibly be achieved by using regions that have been obtained via a region segmentation of the image based on the grey levels. These would in general be non regular regions. This is an area of continuing research.

## 5 Conclusion

Two simple cue rules have been proposed. The first one based on the rate of change of edgel direction. The second is based on a ratio of edgel directions. The first cue can be used to distinguish bushy vegetation from objects that have smooth long lines associated with them (like buildings and cars). The second cue has been shown to roughly discriminate between objects that have long horizontal and vertical lines and others. The cue measures the ratio of the number of horizontal edgels to the number of vertical edgels. The vegetation in the images produces both vertical and horizontal edgels. This is why the cue does not operate well in the presence of vegetation edgels. However, a combination of the vegetation cue followed by linear feature cue improves the performance of the linear feature cue. This has been demonstrated by developing a cue selective between cars and buildings within the test images. This demonstrates one important feature of cues in that the full power is achieved by the application of cues in combination rather

than in isolation.

#### References

- Brisdon, K., 1987, "Alvey MMI-007 Vehicle Exemplar: Evaluation and Verification of Model Instances", Proceedings of the third Alvey Vision Conference, 1987, Univ. of Cambridge, 1987, 15-17 Sept. pp 33- 37.
- Goddon, R.J., Fullwood, J.A., Hyde, J., 1987, "Alvey MMI-007 Vehicle Exemplar: Image Segmentation and Attribute Generation" Proceedings of the third Alvey Vision Conference, 1987, Univ. of Cambridge, 1987, 15-17 Sept. pp 5 - 13.
- Knoll, T.F., and Jain, R.C., 1986, "Recognising Partially Visible Objects Using Feature Indexed Hypothesis", R.A., 2,1, March 1986.
- Lowe, D.G., 1985, "Perceptual Organisation and Visual Recognition", Kluwer, Boston,MA.
- Lowe, D.G., 1987, "Three dimensional object recognition from single two dimensional images", Artificial Intelligence, 31, pp 233-395.
- Morton, S.K., 1987, "Alvey MMI-007 Vehicle Exemplar: Object Hypothesis by Evidential Reasoning", Proceedings of the third Alvey Vision Conference, 1987, Univ. of Cambridge, 1987, 15-17 Sept. pp 15 - 25.
- Sullivan, G.D., "Alvey MMI-007 Vehicle Exemplar: Performance and Limitations", Proceedings of the third Alvey Vision Conference, 1987, Univ. of Cambridge, 1987, 15-17 Sept. pp 39 - 45.

## DOCUMENT CONTROL SHEET

Overall security classification of sheet ..... UNCLASSIFIED

(As far as possible this sheet should contain only unclassified information. If it is necessary to enter classified information, the box concerned must be marked to indicate the classification eg (R) (C) or (S) )

1. DRIC Reference (if known)	2. Originator's Reference MEMO 4109	3. Agency Reference	4. Report Security UNCLASSI- Classification FIED	
5. Originator's Code (if known) 778400	6. Originator (Corporate Author) Name and Location RSRE SAINT ANDREWS ROAD, MALVERN, WORCS WR14 3PS			
5a. Sponsoring Agency's Code (if known)	6a. Sponsoring Agency (Contract Authority) Name and Location			
7. Title GENERIC CUEING IN IMAGE UNDERSTANDING				
7a. Title in Foreign Language (in the case of translations)				
7b. Presented at (for conference papers) Title, place and date of conference				
8. Author 1 Surname, initials FRETWELL P	9(a) Author 2 BAYLISS D A	9(b) Authors 3,4... RADFORD C J	10. Date 1987	pp. ref. 11
11. Contract Number	12. Period	13. Project	14. Other Reference	
15. Distribution statement				
Descriptors (or keywords)				
continue on separate piece of paper				
<p><b>Abstract</b></p> <p>We report an attempt to derive simple low level features which may readily be extracted from an image. The features may be used to provide a classification of objects, so acting as a cue to aid further recognition.</p> <p>It is shown that such an approach may find applications in the early stages of image analysis for the classification of objects in an open world situation. The method is illustrated by application of classification rules based on an estimate of line wigglyness (fractal dimension) and analysis of the directional edgel statistics.</p>				



END

DATE  
FILMED

5 88